U.S. DEPARTMENT OF ENERGY | Office of Science

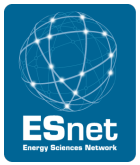
Advanced Scientific Computing Research (ASCR)

ARCHSTONE

(Advanced Resource Computation for Hybrid Service and TOpology NEtworks)

Intelligent Network Services for Advanced Application Workflows

ASCR PI Meeting
Bethesda Maryland
March 1st-2nd, 2012



Personnel

- **USC/ISI**
  - Tom Lehman
  - Xi Yang
- **ESnet**
  - Chin Guok
  - Eric Pouyoul
  - Inder Monga
  - Vangelis Chaniotakis
  - Bharath Ramaprasad (UMass)
- **UNM**
  - Nasir Ghani
  - Feng Gu
  - Kaile Liang

Presentation Outline

- **ARCHSTONE Architecture, Technology, Services Overview**
  - NSI (Network Service Interface)
  - Network Topology and Service Schemas
  - MX-TCE (Multi-Dimensional Topology Computation Engine)
    - Computation Process and Algorithms
  - "Network Service Plane" with "Intelligent Network Services"
    - ask the network "what is possible?" questions
  - Multi-Layer Provisioning (supporting schemas, topology descriptions)
  - Multi-Point Provisioning (supporting schemas, topology descriptions)
- **OSCARS Integration**
  - OSCARSv0.6 extensions to incorporate ARCHSTONE technology
- **Testing and Development Environment**
  - Multi-Layer Provisioning on ANI Testbed
    - ANI Testbed
  - Intelligent Network Services
    - Production Networks

Vision Statement

- **The Next Generation of Advanced Networked Applications (Net+, Cloud based services) will require more "flexible control", "scheduling", and "deterministic performance" across all the resources in their ecosystem**
  - these applications will be user focused and tailored to domain specific requirements
- **This will require integration and co-scheduling across Network, Middleware, and Application level resources**
- **These resources will be heterogeneous in many dimensions (technology, capabilities, policy, and administrative control)**
- **The following will be required to operate in this environment:**
  i. A new class of "Intelligent Network Services" to feed the co-scheduling algorithms and workflows - **ARCHSTONE**
  ii. Co-scheduling algorithms, protocols, and workflows which can operate in this "distributed heterogeneous multi-resource environment"
  iii. Integration of the above items with application specific workflows and systems

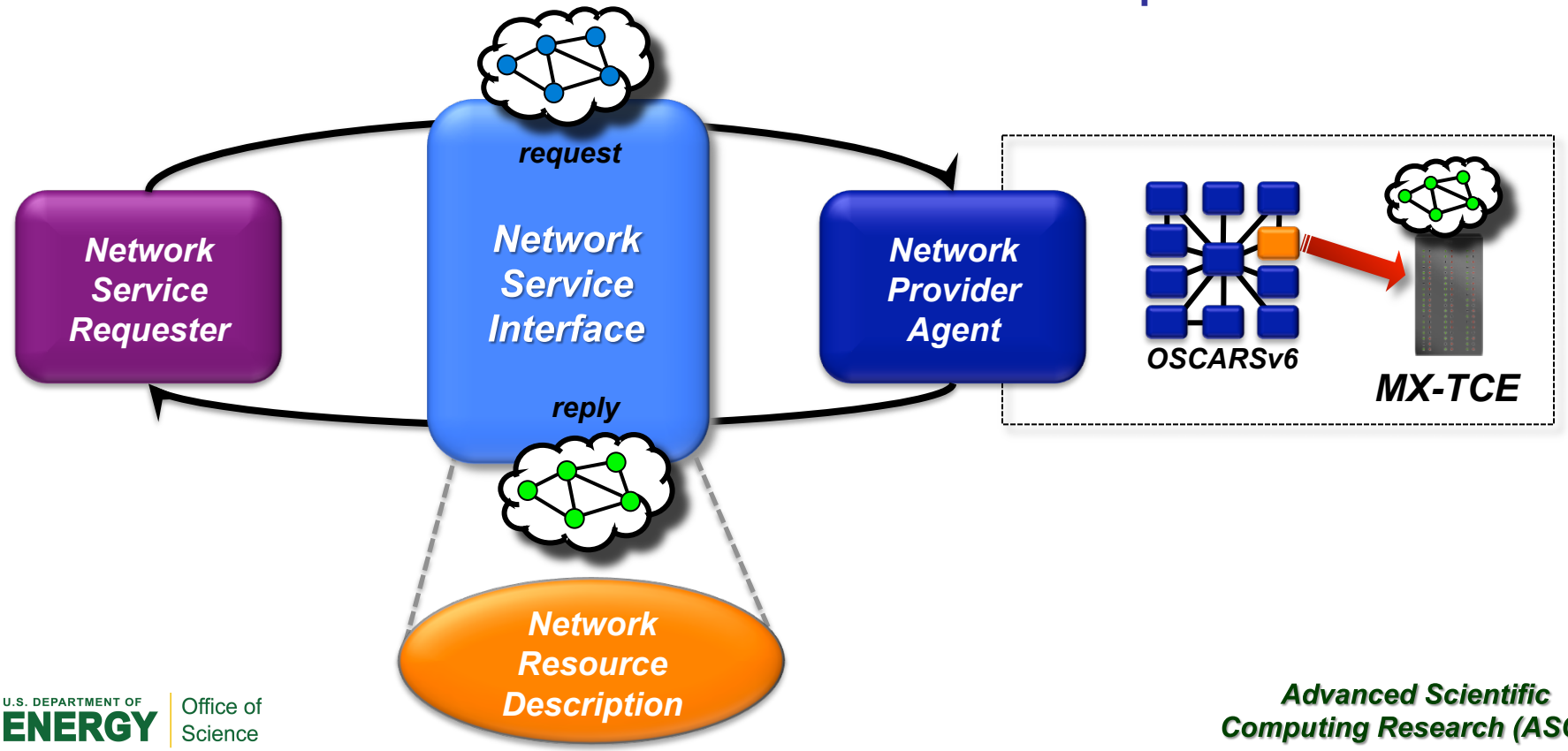
ARCHSTONE

Key Components (added as extensions to OSCARSv0.6)

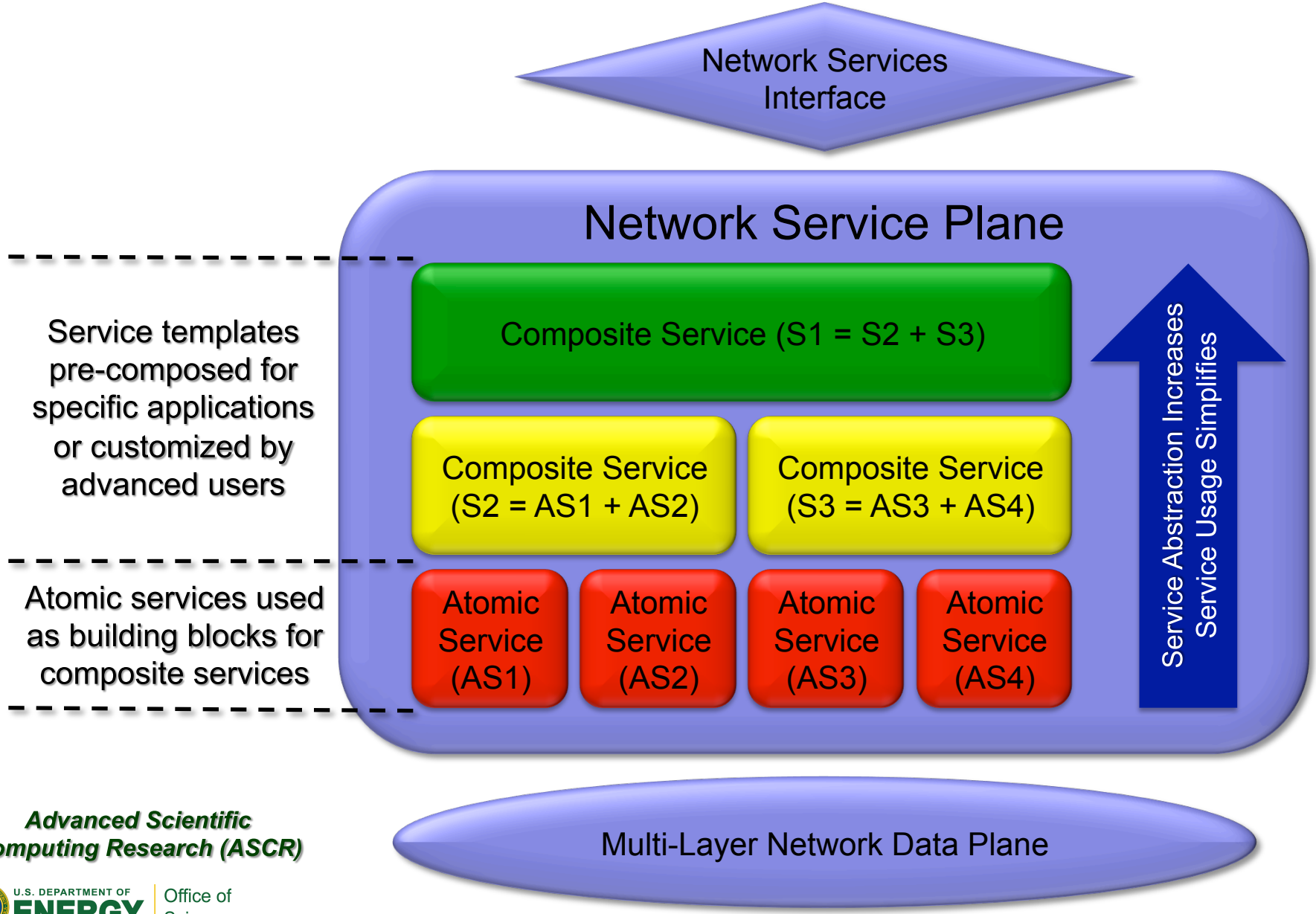
- **Network "Service Plane" formalization**
  - Composable Network Service architecture
  - ARCHSTONE Network Service Interface as client entry point
- **Extensions to Topology and Provisioning Schemas to enable:**
  - multi-layer topologies
  - multi-point topologies
  - requests in the form of a "service-topology"
  - vendor specific features
  - technology specific features
  - node level constraints
- **MX-TCE (Multi-Dimensional Topology Computation Engine)**
  - Computation Process and Algorithms
- **Enable a New class of Network Services referred to as "Intelligent Network Services"**
  - clients can ask the network "what is possible?" questions
  - can ask for "topologies" instead of just point-to-point circuits

ARCHSTONE Architecture Components

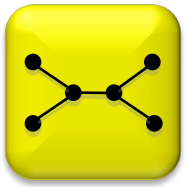
- **Advanced Network Service Interface**
  - "Request Topology" and "Service Topology" concepts
  - Common Network Resource Description schema
  - Network Service Plane access point
- **Multi-Dimensional Topology Computation Element (MX-TCE)**
  - High Performance computation with flexible application of constraints
- **Use OSCARSv6 as base infrastructure and development environment**



Atomic and Composite Network Services Architecture

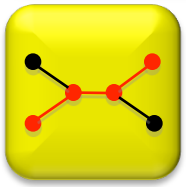


Atomic Services Examples



Topology Service to determine resources and orientation

Resource Computation Service* to determine possible resources based on multi-dimensional constraints ***(*MX-TCE)***



Connection Service to specify data plane connectivity

Protection Service to enable resiliency through redundancy

Restoration Service to facilitate recovery

Security Service (e.g. encryption) to ensure data integrity

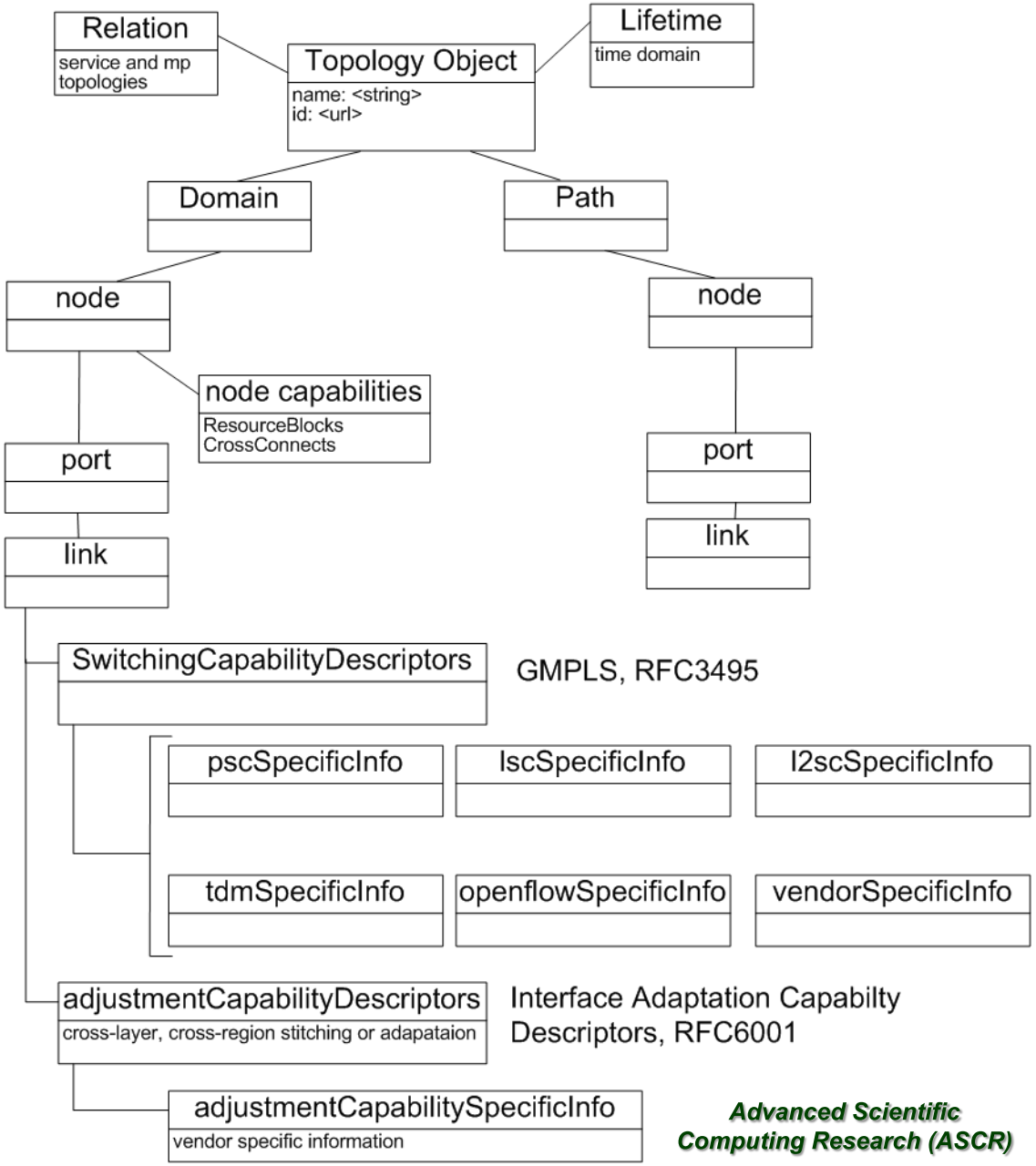
Store and Forward Service to enable caching capability in the network

Measurement Service to enable collection of usage data and performance stats

Monitoring Service to ensure proper support using SOPs for production service

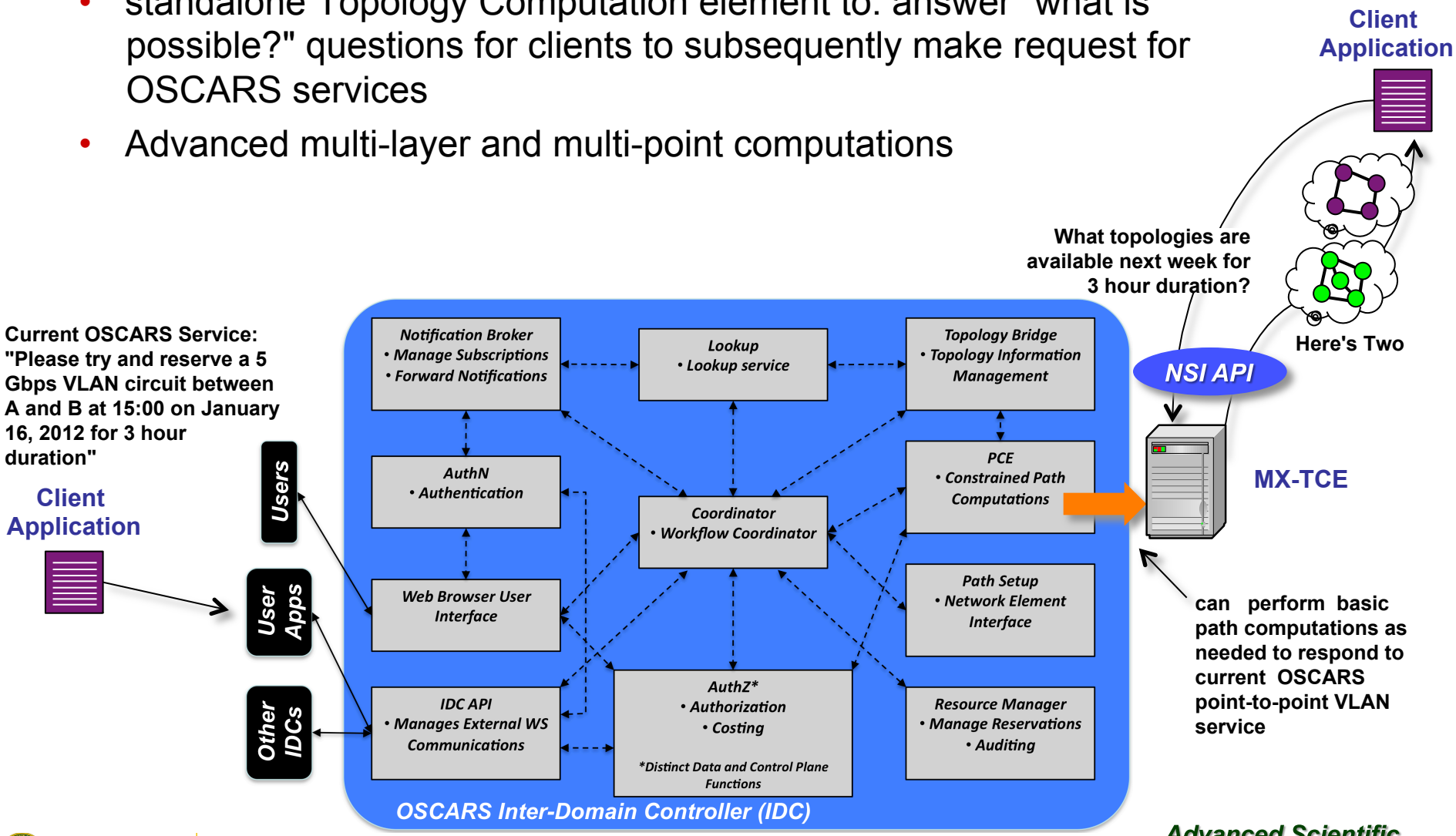
ARCHSTONE Network Schema Extensions

- **Extensions to OSCARS v0.6**
- **Added features for:**
  - multi-layer topologies
  - multi-point topologies
  - requests in the form of a "service-topology"
  - vendor specific features
  - technology specific features
  - node level constraints
- **Result is a schema "Superset" to what OSCARSv0.6 now uses**
  - schema with ARCHSTONE extensions will be backward compatible with current OSCARS operations

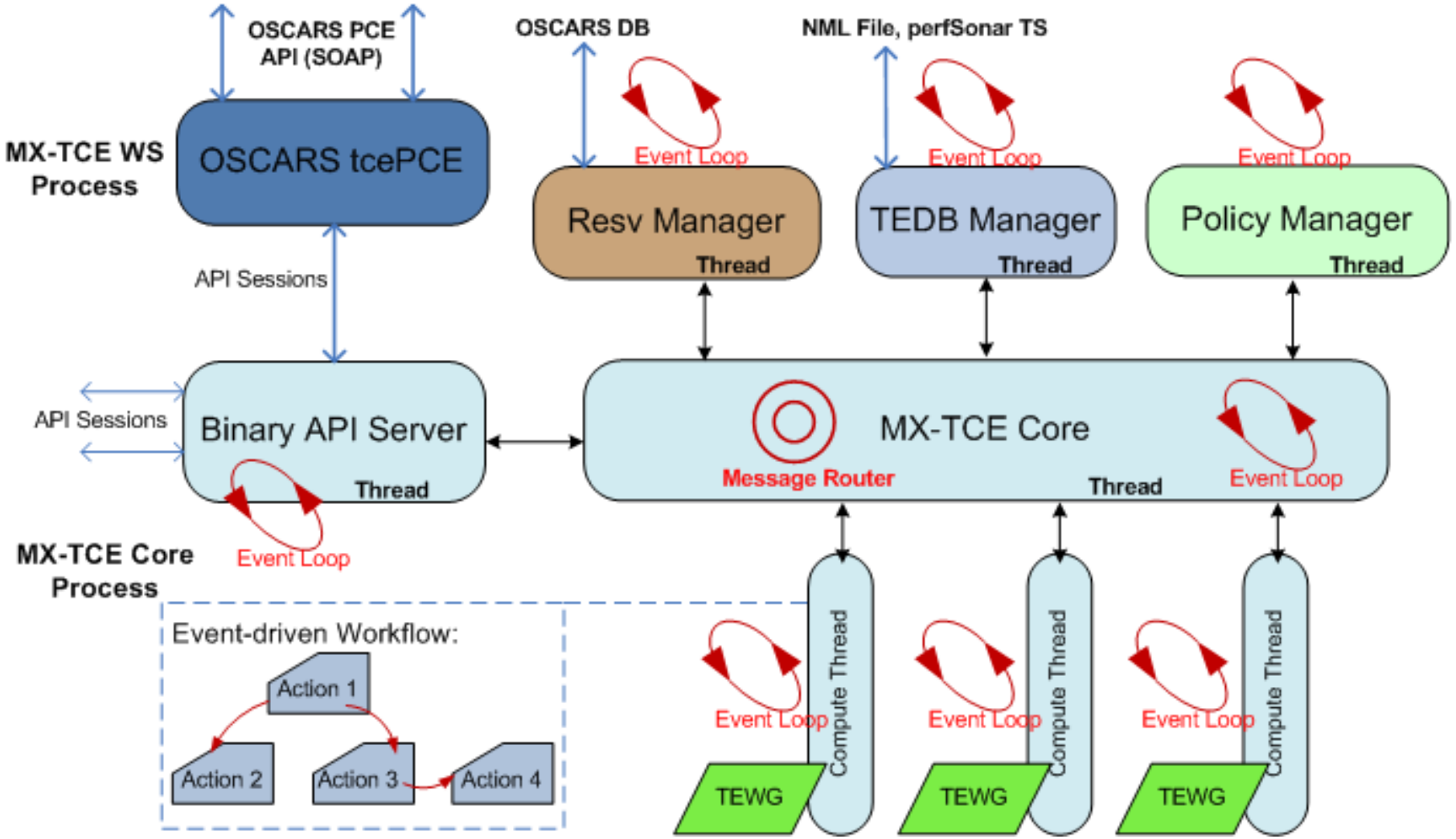


ARCHSTONE Summary

- **MX-TCE role in OSCARS**
  - perform basic path computation for current OSCARS service
  - standalone Topology Computation element to: answer "what is possible?" questions for clients to subsequently make request for OSCARS services
  - Advanced multi-layer and multi-point computations



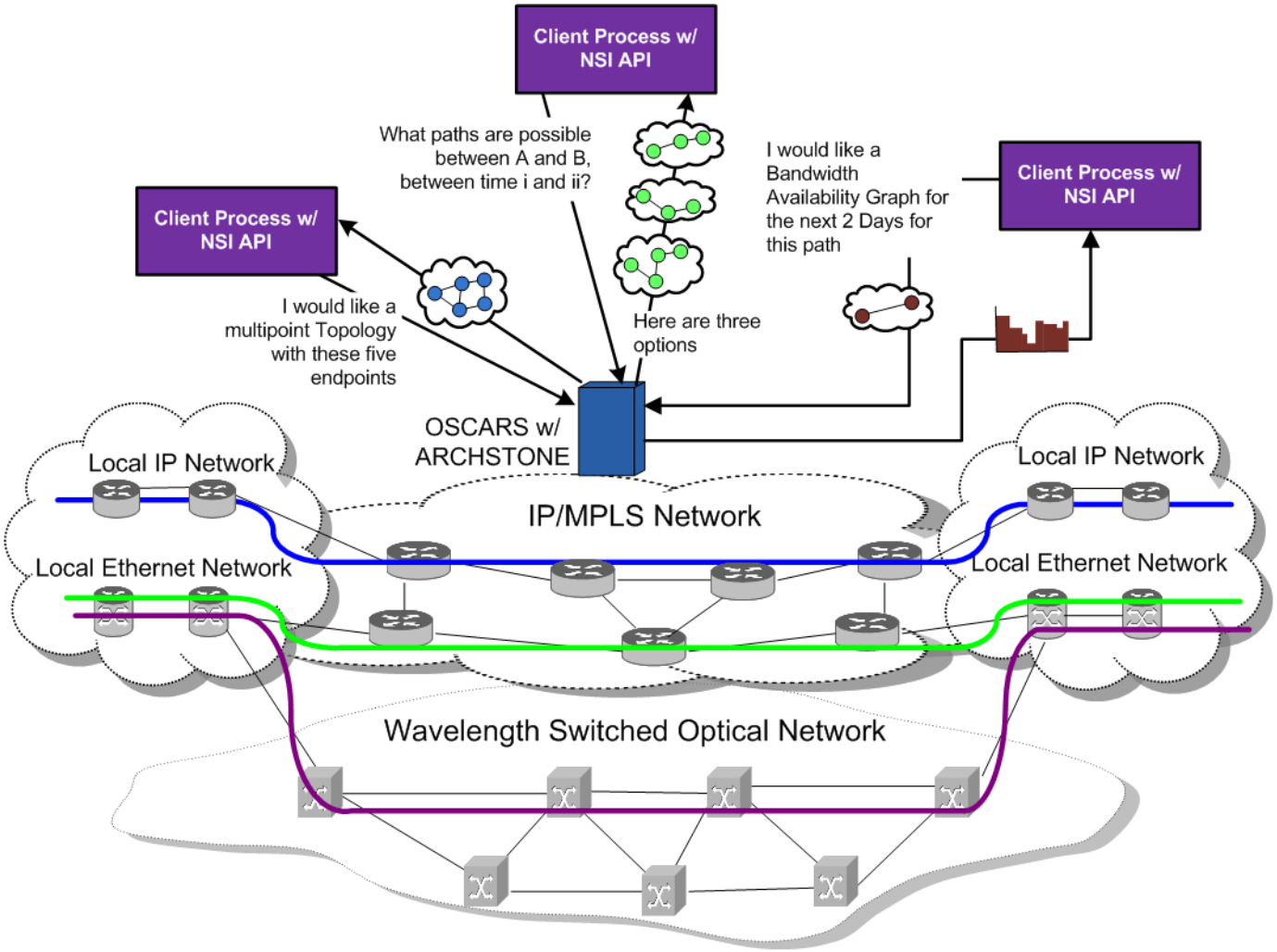
MX-TCE Architecture and Implementation



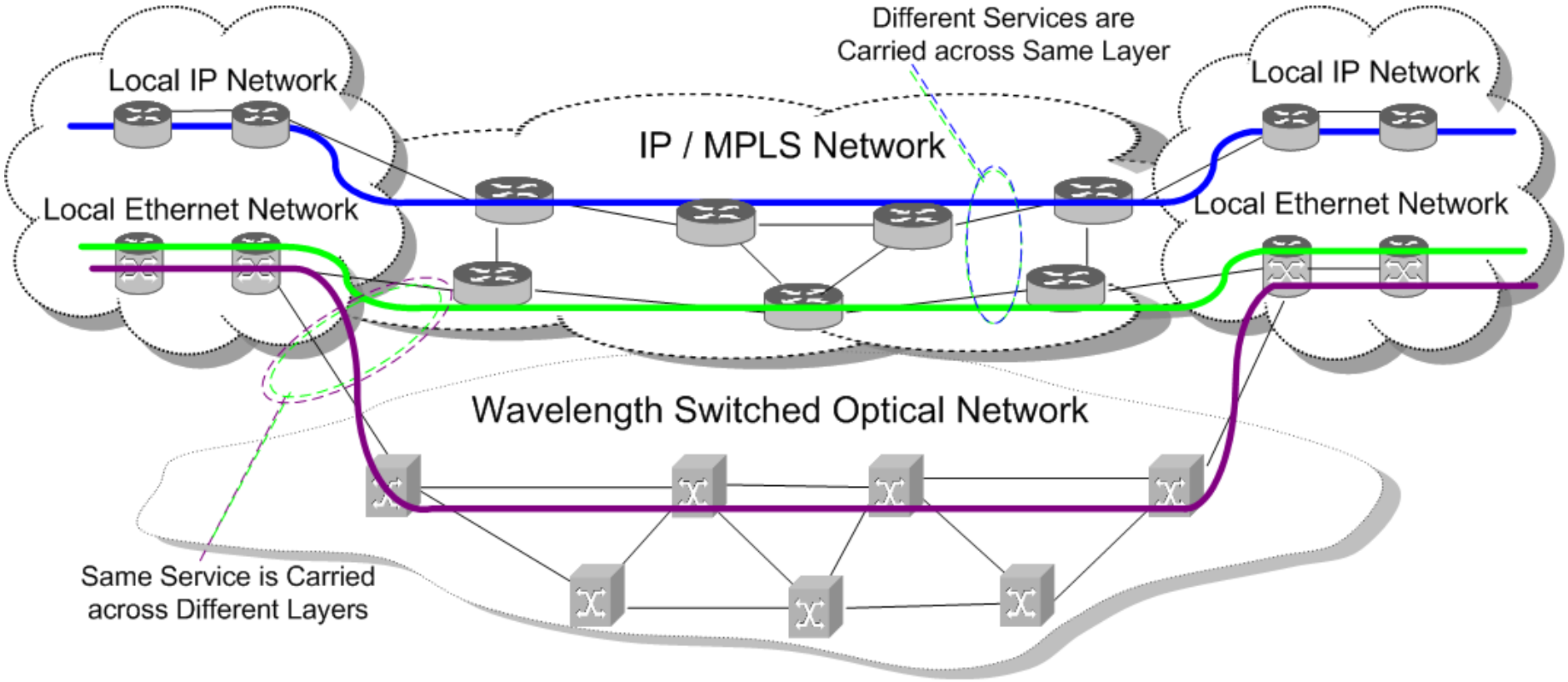
MX-TCE Architecture and Implementation

- **Unified API/NSI support for P2P, Multi-Point, Multi-Layer, schedule and co-scheduling requests under extended NML schema**
- **Implemented OSCARS PCE API to become swappable OSCARS module**
  - support existing OSCARS PCE capability as a single TcePCE
  - support co-scheduling via optionalConstraint extension
- **Multiple path and topology computation workflows**
  - kicked off based on request types: P2P, MP, MLN/MRN, coSheduling etc. and combinations
  - support concurrent requests through multi-threading
- **Modularized differentiated algorithm execution driven by workflows**
- **Transform computation results into**
  - provisioning friendly path object depending on path control scenarios
  - NSI and OSCARS compliant reply messages

Network Service Plane with Intelligent Network Services



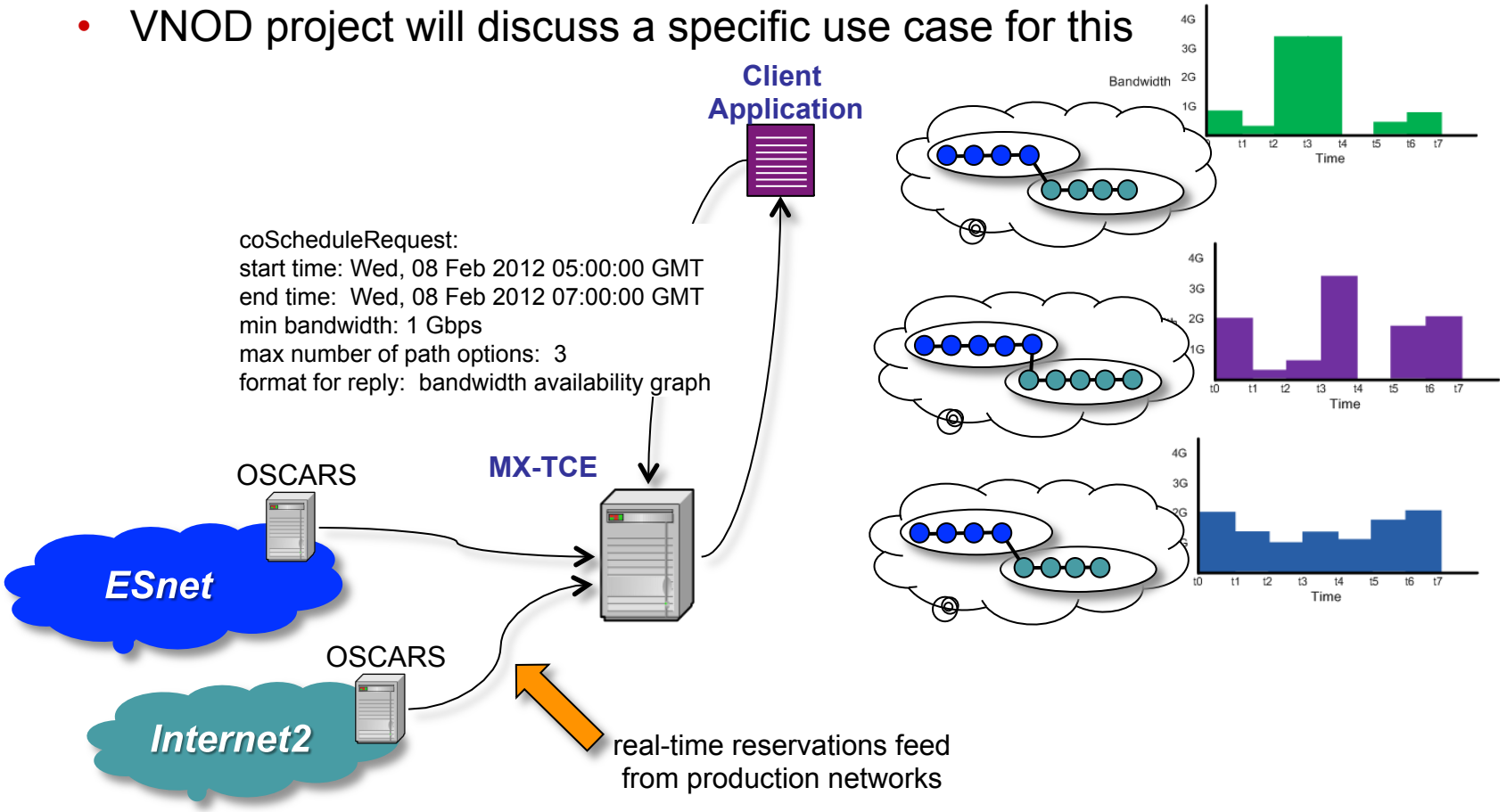
Multi-Layer Services and Provisioning



- **ARCHSTONE Extensions**
  - Multi-Layer Topology Representations
  - Multi-Layer Topology Computations
  - Multi-Layer Provisioning
- **Layer Decisions**
  - driven by resource constraints or client requests for specific performance characteristics (i.e. low latency, low jitter, etc)

Intelligent Network Services Deployment

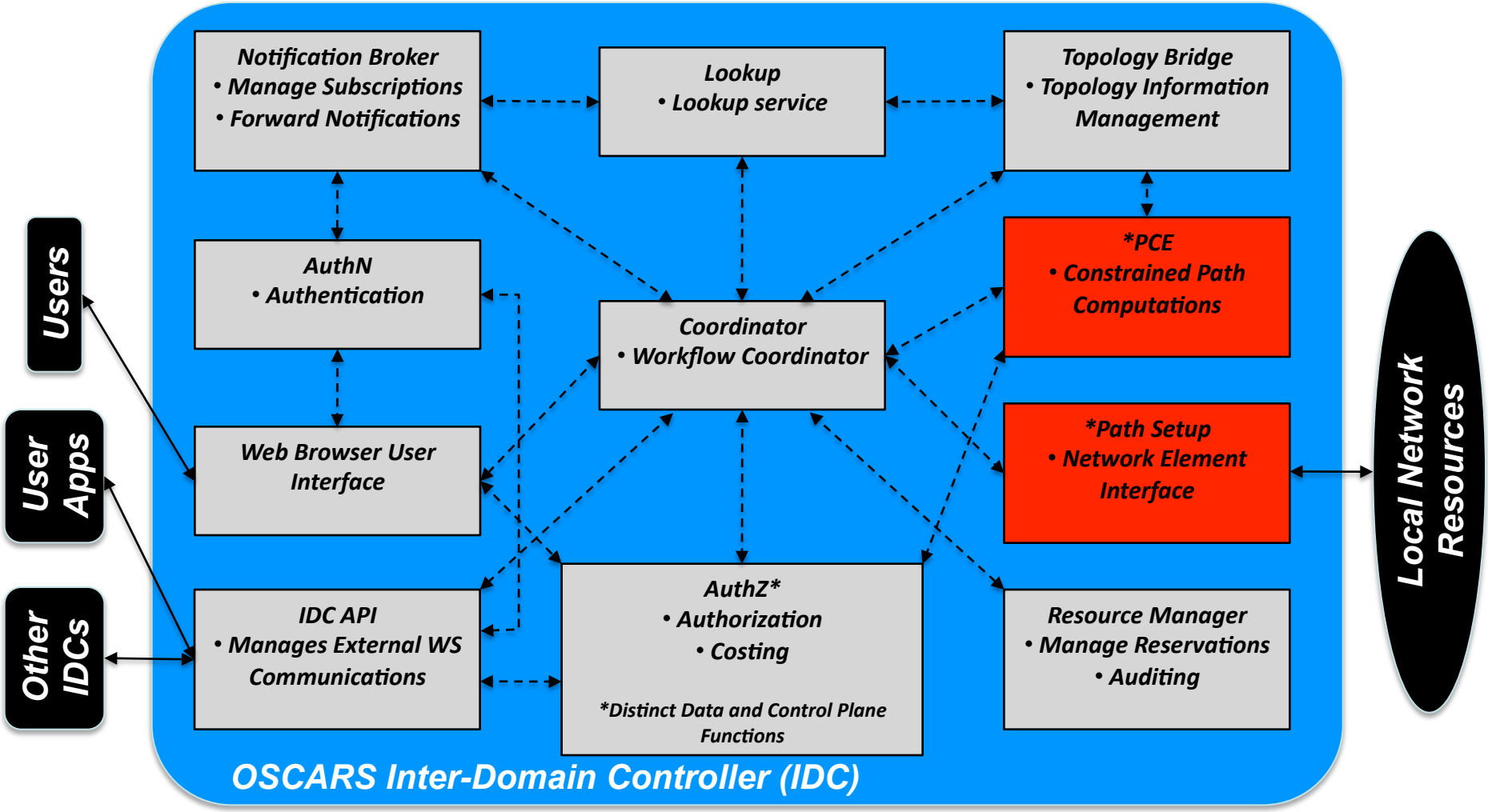
- **Prototype Deployment on Production Networks**
  - ESnet and Internet2
  - Real-time reservations being processed to provide answers to "what is possible?" questions that client can then use to make requests on the operational networks
  - VNOD project will discuss a specific use case for this



Evolution of OSCARS

- **OSCARS started life as a DOE funded research project in 2004 to manage dynamic circuits/bandwidth in the WAN**
- **Up till OSCARS v0.5 the code was tailored specifically to production deployment requirements**
- **In OSCARS v0.6 the entire code base was re-factored to focus on enabling research and production customization**
  - Distinct functions are now individual processes with distinct web-services interfaces
  - Flexible PCE framework architecture to allow “modular” PCEs to be configured into the path computation workflow
  - Extensible PSS module allows for multi-layer, multi-technology, multi-point circuit provisioning
  - Protocol used to make requests to OSCARS (IDC protocol) was modified to include an “optional constrains field” to allow testing of augmented (research) features without disrupting production service model

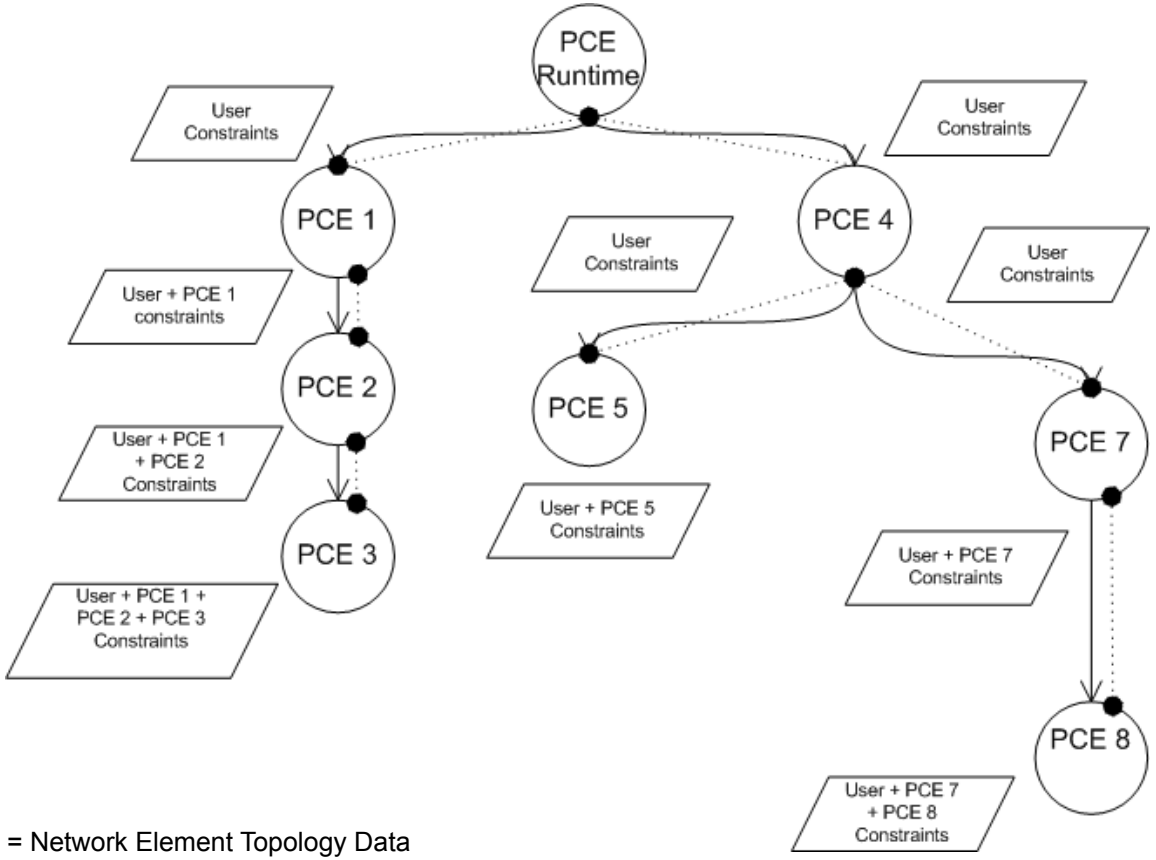
Modularization of OSCARSv0.6



* Current focus of research projects

Flexible PCE Framework

- Supports arbitrary execution of distinct PCEs, e.g.
  - Graph of PCE Modules



Constraints = Network Element Topology Data

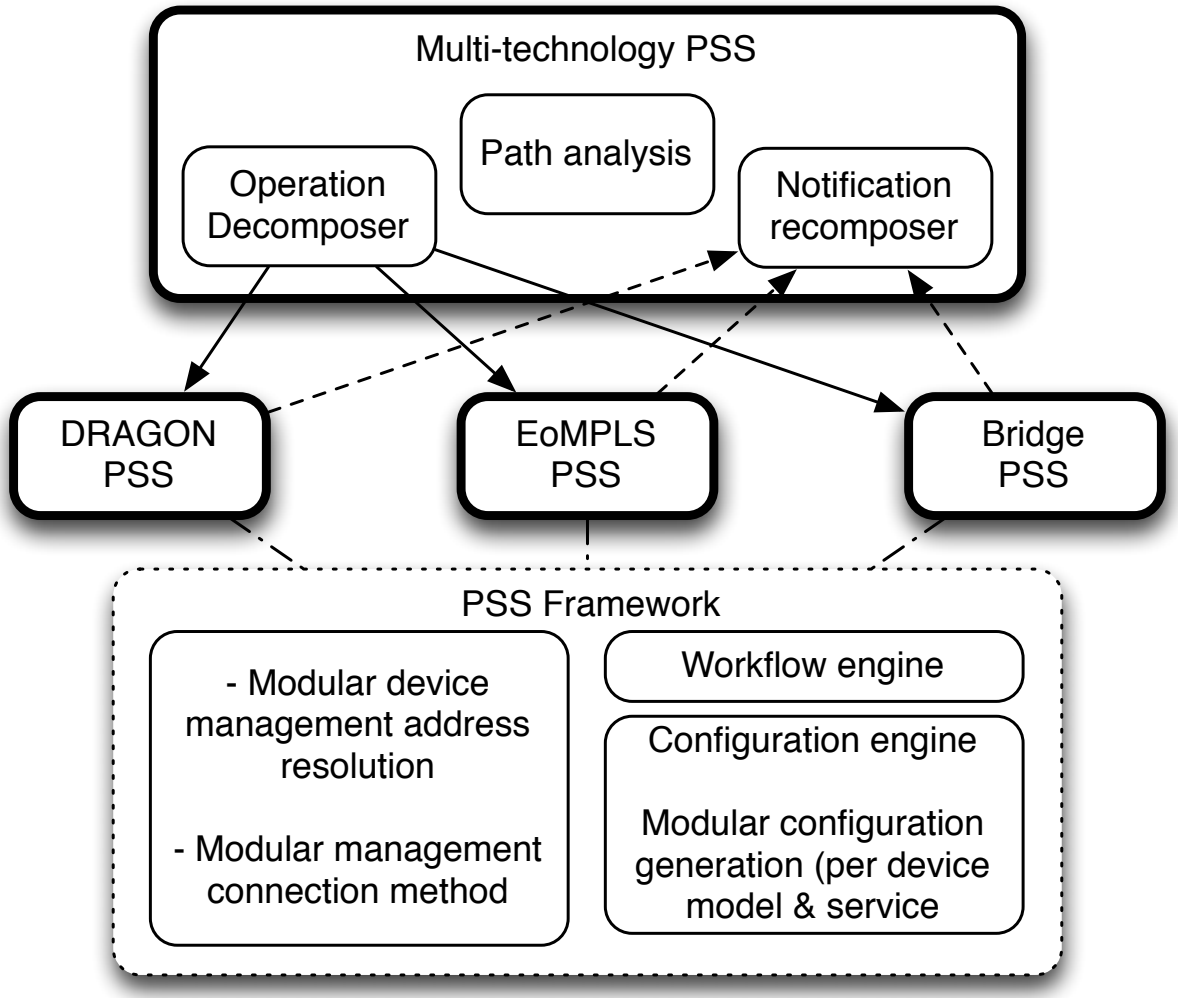
OptionalConstraint Extension to IDCP

- "optionalConstraint" added to support research features without constant need to change base protocol
- Enhancements prototyped in "optionalConstraint" will migrate to base protocol once they have been baked

```
<xsd:complexType name="resCreateContent">
   <xsd:sequence>
      <xsd:element name="messageProperties" type ="authP:messagePropertiesType" maxOccurs="1" minOccurs="0"/>
      <xsd:element name="globalReservationId" type="xsd:string" maxOccurs="1" minOccurs="0"/>
      <xsd:element name="description" type="xsd:string" />
      <xsd:element name="userRequestConstraint" type="tns:userRequestConstraintType" maxOccurs="1"
minOccurs="1" />
      <xsd:element name="reservedConstraint" type="tns:reservedConstraintType" maxOccurs="1" minOccurs="0" />
        <xsd:element name="optionalConstraint" type="tns:optionalConstraintType" maxOccurs="unbounded"
minOccurs="0"/>
    </xsd:sequence>
</xsd:complexType>
...
<xsd:complexType name="optionalConstraintType">
   <xsd:sequence>
      <xsd:element name="value" type="tns:optionalConstraintValue"/>
   </xsd:sequence>
   <xsd:attribute name="category" type="xsd:string" use="required"/>
</xsd:complexType>

<xsd:complexType name="optionalConstraintValue">
   <xsd:sequence >
      <xsd:any maxOccurs="unbounded" namespace="##other" processContents="lax"/>
   </xsd:sequence>
</xsd:complexType>
```

Extensible PSS Module

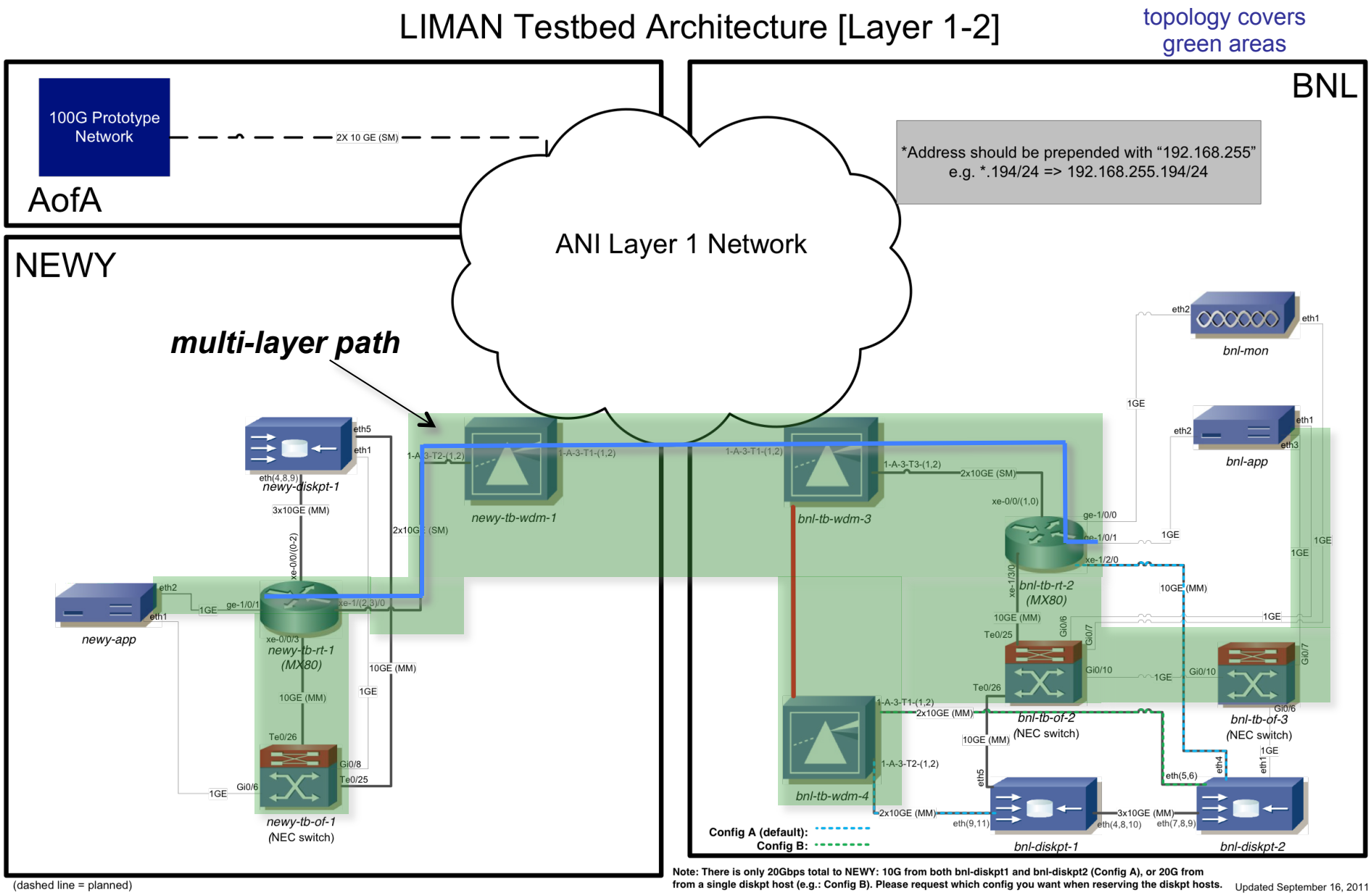


Adoption of OSCARS v0.6

OSCARS v0.6 is starting to gain adoption and see production deployments

- **Field tested at SC11**
  - Deployed by SCinet to manage bandwidth/demo bandwidth on show floor
  - Modified (PSS) by USC/ISI to manage Openflow switches
  - Modified (Coordinator and PSS) by ESnet to broker bandwidth and coordinate workflow
- **Currently deployed in ESnet 100G Prototype Network**
  - Modified (PSS) to support ALU devices and “multi-point” circuits
- **Adopted by Internet2 for NDDI and DYNES**
  - IU GRNOC has modified OSCARS v0.6 (PSS and PCE) to support NDDI OS3E
- **Under review by RNP (Brazilian R&E Network)**
  - RNP has committed to deploying OSCARS (v0.5) in production in 2012, with v0.6 under consideration

Multi-Layer Provisioning Demonstration
ANI Testbed Topology



(dashed line = planned)

Note: There is only 20Gbps total to NEWY: 10G from both bnl-diskpt1 and bnl-diskpt2 (Config A), or 20G from from a single diskpt host (e.g.: Config B). Please request which config you want when reserving the diskpt hosts.

Updated September 16, 2011

Thoughts and Conclusions

- **The architecture adopted by ARCHSTONE, OSCARS:**
  - Centralized at the Intra-Domain level for resource management and service provisioning
  - Distributed at the Inter-Domain level for resource management and service provisioning
  - External topology distribution systems must limit the amount dynamic data exported (scalability and stability issues)
  - Resource identification for real-time service provision can only be done by local domain systems
  - Multi-domain service provision will require chain or tree mode protocols which include real-time negotiation/multi-phase commit features
- **"Intelligent Network Services" is the key capability that needs to be developed next to support co-scheduling across network, middleware, application domains – Network API needed to make service available to workflow engines**
- **OpenFlow/Software Defined Networking offers a set of network capabilities which can enhance these "Intelligent Network Services"**
  - but the "Intelligent Network Services" and co-scheduling technologies are the distinct and key value added feature set that we are addressing

Thank-you

EXTRAS

ARCHSTONE
Additional Information

- **archstone.east.isi.edu**
  - Architecture and Design Documents
  - MX-TCE Software
  - Extensions to OSCARS Topology and Request Schemas
  - Example topology descriptions, service requests, service topologies (responses)
  - ANI Testbed configuration and use
- **OSCARSv0.6 project**
  - code.google.com/p/oscars-idc/
- **OSCARS ARCHSTONE Branch**
  - oscars.es.net/repos/oscars/branches/archstone/

Status and Schedule

- **ARCHSTONE Architecture, Design, (and most) Implementation complete**
  - Schema extensions (multi-layer, multi-point, service topologies, vendor specific technology specific, node level constraints)
  - MX-TCE with capabilities for OSCARSv0.6 service computations
  - Advanced resource computation (multi-layer computations, multi-point computation, "what is possible?" questions)
- **Deployment**
  - Prototype system now deployed on production networks (ESnet and Internet2)
  - Testing underway in collaboration with VNOD project
  - Testing and development continues on the ANI testbed for multi-layer work
- **Immediate to do**
  - Completion of PSS for heterogeneous technology and vendor environments
  - Complete multi-point topology computation and service topology
  - May add a few more "intelligent network service" types based on user requirements
  - modify main OSCARSv0.6 trunk so that it will be backward compatible with new schemas (only a few changes needed)
- **Schedule**
  - complete immediate to do items by spring 2012
  - transition capability to operational networks as a general service available to dynamic network users